Available online at www.sciencedirect.com

ScienceDirect

Remote Sensing of Environment 110 (2007) 412–419

Remote Sensing of Environment

www.elsevier.com/locate/rse

Remote sensing support for national forest inventories

Ronald E. McRoberts [a,*], Erkki O. Tomppo [b]

[a] Northern Research Station, USDA Forest Service, Saint Paul, Minnesota, USA
[b] Finnish Forest Research Institute, Helsinki, Finland

Received 12 August 2005; received in revised form 21 September 2006; accepted 23 September 2006

Abstract

National forest inventory programs are tasked to produce timely and accurate estimates for a wide range of forest resource variables for a variety of users and applications. Time, cost, and precision constraints cause these programs to seek technological innovations that contribute to measurement and estimation efficiencies and that facilitate the production and distribution of an increasing array of inventory data, estimates, and derived products. Many of the recent innovations have involved remotely sensed data and related statistical estimation techniques. Current applications of remote sensing in support of national forest inventories are reviewed for three areas: (1) observation or measurement, meaning using remotely sensed data in lieu of field observations or measurements; (2) estimation, meaning calculation of traditional inventory areal estimates such as forest area or volume per unit area; and (3) mapping. Future applications focus on two areas: augmenting field measurements with remotely sensed data obtained from lidar sensors and Internet accessible map-based estimation.

Keywords: Active sensor; k-nearest neighbors; Stratification; Large area estimation; Mapping

1. Introduction

The mission of a national forest inventory (NFI)[1] is to produce and report timely and accurate estimates of forest resources. The variables for which estimates are produced include, but are not limited to, forest area, volume, condition, growth, mortality, removals, trends, and forest health. Estimates are reported for these variables for categories of forest types or species, ownerships, silvicultural and cutting regimes, and political units such as municipalities, counties, and provinces or states. Users of inventory data are many, including forest land planners and managers, forest industry decision makers, and environmental groups. Increasingly, forest inventory data and estimates are used to satisfy international reporting requirements (e.g., United Nations Food and Agriculture Organization Forest Resource Assessment; United Nations Framework Convention on Climate Change; Land Use, Land Use Change, and Forestry; Kyoto Protocol) and to assess the sustainability of forest management practices in accordance with the criteria and indicators specified by the Ministerial Conference on the Protection of Forests in Europe (MCPFE, 1990) and the Montréal Process (Montréal Conference Working Group, 2005).

Because complete censuses of all trees on all forest lands are prohibitively expensive and time-consuming, NFIs rely on sample-based procedures to produce areal estimates of forest resources such as forest area and volume per unit area. A wide variety of sampling designs have been used, although most now have systematic components that prescribe sampling units on either regularly spaced grids or in regular polygons that tessellate the area of interest (AOI) (McRoberts et al., 2005). The sampling units vary with respect to factors such as size, individual or cluster plots, permanent or temporary plots, and fixed or variable radius plots. Plot observations include, but are not limited to, area of forest cover; land ownership, use, and productivity; tree species, diameter, and height; forest health; biological diversity; and soil attributes.

As the uses and applications for inventory data and estimates increase, so do the number of variables requiring observation or

* Corresponding author. 1992 Folwell Avenue, Saint Paul, Minnesota 55108, USA. Tel.: +1 651 649 5174; fax: +1 651 6495285.
E-mail address: rmcroberts@fs.fed.us (R.E. McRoberts).

[1] In this context, national forest inventory (NFI) refers to an inventory conducted at the national level as per the European use of the term, not an inventory of a national forest as the term might be construed in the United States of America.

0034-4257/$ - see front matter © 2007 Published by Elsevier Inc.
doi:10.1016/j.rse.2006.09.034

measurement. For example, European NFIs typically collect field information on 100–400 variables. As the number of variables increases, so do the complexity, cost, and time necessary to conduct the inventories. Thus, NFIs seek technological improvements for increasing the speed and cost efficiency of conducting the inventories while simultaneously increasing the precision and timeliness of an ever widening array of estimates. The advent of low cost, widely available, remotely sensed data has been the basis for many of the important recent technological improvements. Remotely sensed data have not only contributed to increasing the speed, cost efficiency, precision, and timeliness associated with inventories, but they have facilitated construction of maps of forest attributes with spatial resolutions and accuracies that were not feasible even a few years ago.

The primary objectives of the discussion are to describe the methods by which remote sensing contributes support to modern NFIs. The discussion focuses on three applications of remotely sensed data: (1) surrogates for field observation or measurement; (2) ancillary data to improve the precision of traditional inventory areal estimates; and (3) mapping.

2. Observation or measurement

Remotely sensed data may be used in lieu of more expensive ground observations and measurements. Aerial photography has been a traditional source of such data, and its increasing availability in digital formats has greatly facilitated its uses. For many NFIs, the cost of travel to and from a plot location is the greatest portion of the expense of plot measurement. In areas with sparse forest or a mixture of agricultural and forest land uses, such as the prairie and plains region of the United States of America (USA), considerable cost savings may be realized by initially observing plot locations using aerial photography or high resolution satellite imagery. The cost of acquiring and interpreting the remotely sensed data for all plots is much less than the cost of traveling to and from the large subset of plots having no forest cover or forest land use. When photography is available in a digital format even greater cost savings may be realized. Integrating photography in digital formats with other spatial data in geographic information systems permits rapid and easy identification of navigational impediments such as land forms or swamps and measurement of distances to natural features such as water and man-made features such as roads and buildings.

With the advent of active laser and radar sensors, techniques for obtaining vertical forest structure information from remotely sensed data are approaching operational feasibility. Inventory applications for scanning light detection and ranging (lidar) systems include volume and biomass estimation using both large- and small-footprint systems. Large-footprint lidar systems produce estimates of mean height, canopy cover, or canopy density for an area. Regression models constructed using ground measurements, the lidar data, and ancillary data may then be used to predict volume or biomass for stands (Holmgren et al., 2003; Lefsky et al., 1999; Means et al., 1999; Næsset, 1997; Nilsson, 1996; Wallerman & Holmgren, 2007-this issue). Small-footprint lidar systems produce crown width and height measurements for individual trees. Species-specific regression models may then be used to predict tree diameter, usually diameter at breast height (dbh). Height and diameter are then used as inputs to additional species-specific regression models for predicting volume, and/or biomass (Parker & Evans, 2004; Parker & Glass, 2004). Although predictions of diameter and volume may have considerable uncertainty for individual trees, if the models are unbiased, then stand- or plot-level estimates of mean diameter of trees or total volume of growing stock may still be acceptably precise. For coniferous plots in the western USA with basal areas ranging from 0 to 90 m^2/ha, Lefsky et al. (1999) and Means et al. (1999) reported $R^2 > 0.85$ for relationships between observed and estimated total above ground biomass and total basal area.

Radar systems are also under investigation for similar applications. Fransson et al. (2000) used backscattering amplitude data from a synthetic aperture radar (SAR) sensor as inputs to linear regression models for estimating stand-level stem volume, stem diameter, and tree height. Holström and Fransson (2003) reported that the combination of optical and SAR data produced areal estimates of stem volume, age, and proportion of conifers that were superior to those obtained using only a single sensor. Kellndorfer et al. (2004) used digital elevation data collected from the 2000 Shuttle Radar Topography Mission (SRTM) with regression models to estimate vertical vegetation structure. Woodhouse and Hoekman (2000) used tree growth models to obtain information for polarimetric airborne SAR scattering components for boreal-type forests. Data from new polarimetric SAR satellite instruments are expected to provide additional opportunities for using space borne microwave remote sensing in forestry applications; examples of these sensors include the Phased Array type L-band Synthetic Aperture Radar (PALSAR) of the Advanced Land Observing Satellite (ALOS) (www.nasda.go.jp/projects/sat/alos/index_e.html) and RADARSAT 2 (www.radarsat2.info/).

Many studies reporting use of data from active sensors are limited in either spatial extent or degree of complexity of forest conditions. Operational implementation for an NFI requires comparable results for closed canopy and mixed species forests and requires estimation methods for understory trees, also. Regardless of the quality of estimates obtained using data from active sensors, remote sensing measurement of plots is not expected to replace completely field measurement any time in the near future. Nevertheless, approaches based on techniques such as double sampling for regression using lidar measured plots in the first phase as described by Parker and Evans (2004) merit consideration.

3. Areal estimation

3.1. k-Nearest Neighbors

Numerous estimation and mapping approaches have been successfully used by NFIs. However, in recent years, nearest neighbor techniques have received considerable attention and merit special discussion. This discussion focuses on the use of the k-nearest neighbor (k-NN) technique with satellite imagery, but there are other variations such as most similar neighbor (MSN) (Hassani et al., 2004; Moeur & Stage, 1995) and gradient nearest neighbor (GNN) (Ohmann & Gregory, 2002). Further, all variations can be used equally well with sources of ancillary information other than satellite imagery.

Pioneered for forest inventory purposes by the Finnish NFI (Tomppo, 1991, 1996; Tomppo & Halme, 2004), the k-NN technique is a non-parametric, multivariate approach to imputing observations or combinations of observations from sampling units to estimation or mapping units. For an AOI, the set of all pixels for which predictions of inventory variables are desired is denoted the target set, and the set of all pixels containing centers of sampling units or plots as is denoted the reference set. The plots are assumed to be of adequate size to describe the pixels containing their centers. Thus, the elements of the reference set are equivalently characterized as either pixels or as plots. For each pixel, p, in the target set, $I=\{i_1(p),\ldots,i_k(p)\}$ denotes the set of k plots in the reference set nearest to p in the covariate or feature space with respect to a distance metric, d. Common selections for the distance metric are Mahalanobis distance (Kendall & Stewart, 1968) or weighted Euclidean distance,

$$d_{i,p} = \sqrt{\sum_{l=1}^{L} v_l (x_{i,l} - x_{p,l})^2}, \tag{1}$$

where $d_{i,p}$ denotes the distance in feature space between pixels i and p; l indexes the feature space covariates; and the set $\{v_1\}$ consists of weights associated with individual covariates. The k-NN prediction for pixel, p, is,

$$\hat{y}_p = \sum_{i \in I} w_{i,p}\, y_i, \tag{2}$$

where y_i is the vector of observations for the ith plot in the reference set, and $\hat{y}_p$ is the vector of imputed or predicted values for pixel p. The weight, $w_{i,p}$, of plot i to target pixel p is,

$$w_{i,p} = \begin{cases} \dfrac{d_{i,p}^{-t}}{\sum\limits_{i \in I} d_{i,p}^{-t}} & i \in I \\ 0 & \text{otherwise} \end{cases}, \tag{3}$$

where $t\,\varepsilon\,[0,2]$. Common selections for t include $t=0$, which weights all reference set plots equally, and $t=1$ or $t=2$ which weight plots inversely to their feature space distance or distance squared from pixel p. When $t=0$, the k-NN prediction reduces to,

$$\hat{y}_p = \frac{1}{k}\sum_{i \in I} y_i.$$

3.2. Expansion factors

When using plot observations and measurements as the basis for estimates of inventory variables for larger AOIs (e.g., municipalities, counties, provinces or states), the per unit area observation for each plot must be multiplied or expanded by the area the plot represents to obtain an estimate of the total for the AOI. For example, if 75 trees are observed on a 0.1 ha plot, then the observation converts to a per unit area observation of $\frac{75\ \text{trees}}{0.1\ \text{ha}} = 750\ \text{trees/ha}$. If the sampling design features, for example, one plot per 2000 ha, then one possibility for the plot expansion factor is 2000 ha, in which case the total trees represented by the plot observation is 1.5 million trees. More accurate large area estimates may be obtained when plot expansion factors are derived from landscape features. For example, if the plot whose observation of 750 trees/ha is the only plot on a particular soil type and there are 2500 ha of this soil type in the AOI, then a plot expansion factor of 2500 ha would contribute toward a better estimate of total trees on the landscape than would the 2000 ha expansion factor. This approach may be used to obtain expansion factors for both fixed and variable area plots. In the USA, where strata for stratified estimation are derived from classified satellite imagery, expansion factors are calculated as the ratios of strata sizes and the numbers of plots assigned to strata.

Remotely sensed data may contribute substantially to increasing the quality of expansion factors and the precision of estimates. The Finnish NFI uses an innovative application of the k-NN technique in which the expansion factor for the jth plot in the uth AOI is calculated as,

$$c_{j,u} = \sum_{p \in U} w_{j,p}, \tag{4}$$

where p denotes a pixel in the AOI, and $w_{j,p}$ is calculated using Eq. (3). In essence, the plot expansion factor is a function of the number of pixels in the AOI that are close in feature space to the pixel containing the plot center (Tomppo, 1996). Feature space and the corresponding distance metric need not be restricted to satellite image data but may include other ancillary data (Tomppo & Halme, 2004).

3.3. Stratified estimation

Due to budgetary constraints and natural variability, sufficient numbers of plots frequently cannot be measured to satisfy precision guidelines for the estimates of many inventory variables unless the estimation process is enhanced using ancillary data. Classified satellite imagery has been accepted as a source of ancillary data that can be used with stratified estimation techniques to increase the precision of estimates with little corresponding increase in costs. The genesis of this approach is the statistical technique double sampling for stratification as used by NFIs with aerial photography. In the first phase, an initial, relatively dense, systematic sample is observed using aerial photography. The primary purpose of the first phase is to assign these sample units or photo plots to classes, often with respect to the size, density, and species of trees observed on the photography. Information from these classes, sometimes augmented with ancillary data such as soil, land use, and other maps, is used to define strata. The relative areal extent of each stratum is estimated by the proportion of first phase photo plots assigned to the stratum. The second phase consists of field measurement of a subset of the first phase plots. The data for the two phases are combined using stratified estimation formulae (Cochran, 1977). The effectiveness of a stratification is often quantified using relative efficiency, RE, defined as,

$$\mathrm{RE} = \frac{\mathrm{Var}(\overline{Y}_{\mathrm{SRS}})}{\mathrm{Var}(\overline{Y}_{\mathrm{Str}})}, \tag{5}$$

where $\mathrm{Var}(\overline{Y}_{\mathrm{SRS}})$ is the variance obtained under the assumption of simple random sampling (SRS) and no stratification, and $\mathrm{Var}(\overline{Y}_{\mathrm{Str}})$is the variance obtained using stratified estimation.

Poso et al. (1984) and Poso et al. (1987), using the double sampling for stratification technique, obtained the first phase measurements from satellite imagery and defined strata on the basis of classifications of the imagery. Muinonen and Tokola (1990) considered the satellite image information to constitute complete coverage of the AOI rather than just a first phase sample and obtained RE=7.20 for growing stock volume in Finland. Nilsson et al. (2003) used a segmentation approach for defining image-based strata and obtained RE=3.72 for a national estimate of total timber volume in Sweden. McRoberts et al. (2002b) derived strata from the National Landcover Dataset, (NLCD) (Homer et al., 2004; Vogelmann et al., 2001), a 21-class land cover map of the conterminous USA based on Landsat Thematic Mapper (TM) imagery and other ancillary data. The NLCD forest classes were aggregated into one stratum, the non-forest classes into a second stratum, and two additional edge strata were constructed, one on either side of the forest/non-forest boundary. For forest area in the north central USA, they obtained $2.00<\mathrm{RE}<3.50$, and for volume they obtained $1.25<\mathrm{RE}<1.75$. Hoppus and Lister (2003) began with a forest/non-forest classification, reclassified the center pixel in each 5×5 pixel block into one of 26 classes, depending on the number of pixels classified as forest in the 5×5 pixel block, and then aggregated the 26 classes into a smaller number of strata. For estimates of forest area in the northeastern USA, they obtained $1.69<\mathrm{RE}<2.12$. McRoberts et al. (2006) proposed an approach to stratification in which a logistic regression model is used to predict proportion forest area for each pixel. Strata are defined in terms of categories of predicted proportions, $\hat{p}$ (e.g., $0.0\leq\hat{p}\leq0.1$, $0.1<\hat{p}\leq0.5$, $0.5<\hat{p}\leq0.9$, and $0.9<\hat{p}\leq1.0$). For forest area and volume in the north central USA, RE=5.87 and RE=2.71, respectively, were obtained.

Some precautions are necessary when strata are derived using data from plots that are to be stratified, because an assumption underlying stratified estimation is that the plots assigned to a stratum are a random sample of the stratum. In particular, this assumption must be carefully considered when using the k-NN method with a small k-value to obtain predictions from which strata will be derived. The concern is that for small k-values, the set of plots assigned to each stratum will be very similar to the mathematical union of the sets of k-nearest neighbors used to obtain predictions for the pixels assigned to the stratum. The result is that the plots assigned to a stratum may not be a random sample of the stratum. Two approaches to circumvent this problem may be considered. First, the AOI may be subdivided into mutually exclusive subareas, and the plots in one subarea may be used as the reference set for calculating predictions for the other subarea. Second, large values of k may be considered for k-NN prediction. The second approach is supported by the finding of Breidt and Opsomer (2004) that for regression-based predictions, even relatively small calibration data sets circumvent the problem.

3.4. Efficiencies

The contribution of remotely sensed data to increasing areal estimation efficiency is illustrated with Finnish and American examples. For eight AOIs in eastern Finland, each of approximately 10,000 ha, approximately 500 plots per AOI would be necessary to obtain a standard error of 5% for the estimate of mean volume of growing stock. However, using satellite imagery to calculate expansion factors using Eq. (5), a sampling intensity of only one plot per 288 ha (35 plots per 10,000 ha) produced a standard error of the estimate of mean of approximately 5% (Katila & Tomppo, 2001; Katila, 2006). Thus, given that remote sensing costs for the Finnish NFI are only approximately 5% of total inventory costs, this greater than 10-fold reduction in sampling intensity means that remotely sensed data contributes substantially to increasing inventory efficiency.

For the American example, REs obtained using stratified estimation may be translated into cost savings. For two large study areas in the State of Minnesota, USA (21.8 million ha) McRoberts et al. (2002a) reported RE=5.59 and RE=5.38 for stratified estimation of forest area when using strata derived from classified satellite imagery. RE calculated using Eq. (5) may be considered the factor by which a sample size must be increased to achieve the same precision under the assumption of simple random sampling (i.e., not using the remotely sensed data) as was achieved using stratified estimation (i.e., using remotely sensed data). Cost savings were calculated by applying the mean, RE=5.49, to the entire State under four assumptions: (1) the State is approximately 38% forested, (2) only forested plots are visited by field crews, (3) the cost of measuring a plot is approximately US $800, and (4) the sampling intensity is approximately one plot per 2400 ha. For these conditions, the costs savings were US $12.4 million when using stratified estimation with strata derived from classified satellite imagery.

4. Map products

4.1. Answering the “Where?” question

Traditionally, NFIs have used data collected from field plots to respond to the user question “How much?” by reporting plot-based estimates of forest resources for municipalities, counties, and provinces or states. Increasingly, however, users are also asking “Where?” and are requesting that NFIs report not only tabular estimates but also produce maps depicting the spatial distribution of forest resources. Thus, NFIs have initiated efforts to construct maps of forest resources, often using satellite imagery of at least moderate resolution.

In addition, maps based on inventory data may be used to investigate potential sampling designs. The Finnish NFI, which uses a combination of permanent and temporary plots, has used border-to-border forest attribute maps based on satellite imagery and previous NFI data to select sampling designs since the 1990s (Henttonen, 1991; Tomppo et al., 2001). For each potential design, simulated samples of size 1000–2000 are drawn from the maps, and estimates and standard errors are calculated for variables such as growing stock volume per unit area, volume per unit area by species, and proportion forest by species. By combining travel and measurement costs for each design with simulated precision estimates, potential designs may be compared and an appropriate selection may be made. This approach was used for the entire Finnish NFI9 (1996–2003). In addition, in northern Finland, a land cover map constructed from multi-source inventory data using the k-NN technique was used to derive a stratified sampling design and formed the basis for post-stratification in the estimation phase. The land cover classes correspond to potential land use classes such as forest land, poorly productive forest land, or permanently treeless areas.

Inventory mapping applications generally are expected to satisfy several requirements simultaneously. First, the mapping approach should be multivariate because compatibility among the maps of related variables is essential. Separate, univariate maps of variables such as forest land and volume per unit area inevitably depict high volume for some pixels that are also depicted to have little or no forest land. Second, the mapping approach must accommodate the varied forms of the distributions of inventory variables. For example, most parametric multivariate techniques require that the suite of variables have a multivariate Gaussian distribution. Third, the approach should be versatile to facilitate construction of similar and comparable maps for diverse and geographically separated AOIs using different types and resolutions of satellite imagery. The k-NN method seems ready-made for NFI mapping applications. The k-NN method is multivariate, non-parametric, intuitive, and easy to implement. Because it is non-parametric, no assumptions regarding the distributions of variables are required. In addition, for small values of k, k-NN predictions preserve much of the correlation structure among observations of inventory variables. Further, subject to a few requirements such as the plot observation being an adequate sample of the image pixel, the k-NN technique can be readily implemented using a variety of reference data and resolutions of imagery. Finally, as the Finnish NFI has demonstrated, ancillary data such as soil and land use maps may be used to enhance the quality of k-NN predictions; e.g., by avoiding prediction of positive volumes predictions on non-forest land (Katila & Tomppo, 2001).

The k-NN technique has been used extensively to map forest attributes: volume in Finland (Tokola et al., 1996; Tomppo, 1996); forest cover type and basal area in the USA (Franco-Lopez et al., 2001); proportion forest in the USA (McRoberts et al., 2002a); age, height, and volume in Sweden (Reese et al., 2003); age and volume in Sweden (Holmström & Fransson, 2003); forest cover type in Austria (Koukal et al., 2005); species groups in Norway (Gjertsen, 2007-this issue); volume, stocking, diameter, height, and basal area in Ireland (McInerney et al., 2005); and forest area, basal area, volume, and stem density in the USA (McRoberts et al., 2007-this issue). Such maps are used operationally in forest management planning and industrial timber procurement planning (Tomppo, 1996), as well as ecological applications (Pakkala et al., 2002). The georeferenced nature of the maps facilitates integrating data such as a growing stock map with other georeferenced data such as maps of soil classes, site fertility, cumulative growing season temperature, and meteorological information maps. For example, growing stock and soil fertility maps may be integrated to predict forest productivity and to analyze potential forest production scenarios for areas of arbitrary sizes as ordered by clients. The Finnish NFI has already used this approach to develop and analyze harvest scenarios at the municipality level. For this approach, sufficiently accurate increment predictions for individual mapping units or small aggregations of mapping units are necessary.

4.2. Accuracy assessment of map products

Accuracy assessments of forest attribute maps have two components, bias and precision. For most reported k-NN applications, bias is not a severe inhibiting factor at either the pixel or multiple pixel levels (e.g., Katila & Tomppo, 2001; Tomppo & Halme, 2004). The precision of maps may be evaluated at both the individual and multiple pixel levels. The precision of pixel level predictions have been widely reported (e.g., Franco-Lopez et al., 2001; Katila & Tomppo, 2001; Tokola et al., 1996; Tomppo & Halme, 2004) with coefficients of variation usually high, often in the range 0.65–0.80 for mean growing stock volume. However, the precision can be increased and bias can be decreased using techniques reported by Halme and Tomppo (2001) and Tomppo and Halme (2004).

Although the classifications of individual pixels may be relatively poor, the distribution of pixel classifications may still be approximately correct. For example, using ground plot observations, Franco-Lopez et al. (2001) found that the proportions of individual pixels correctly classified with respect to cover types ranged from 0.58 to 0.69, but that the proportions of pixels classified into each cover type for a large AOI were approximately correct. Similarly, Tokola (2000) reported that estimates for volume were always more accurate for multiple pixel AOIs than for individual pixels when using sample plot data for comparisons. The phenomenon has also been reported by Katila and Tomppo (2001), and Tomppo et al. (2002). McRoberts et al. (2007-this issue) found forest area, basal area, volume, and stem density estimates obtained by aggregating pixel-level predictions for circular AOIs of radius 10 km to be not statistically significantly different than estimates obtained directly from inventory plot observations. When pixel prediction errors for continuous variables are independent and distributed with zero mean, and when the number of pixels in the AOI increases, then the variance of the estimate for the AOI converges to zero on the basis of the law of large numbers.

Estimating the precision of estimates for multiple-pixel AOIs obtained by aggregating the pixel predictions is a difficult task because of spatial correlation among the predictions and residuals. Some discussion of this topic is provided in the next section.

4.3. Map-based estimation

Frequently, the tabular estimates of forest inventory variables provided by NFIs are not adequate to satisfy all user requirements; for example, AOIs based on ecological rather than political boundaries or fragmented AOIs. For these kinds of analyses, users may request direct access to inventory data so they can conduct their own analyses. When the requests do not require exact plot locations, there are few constraints on data access. However, if exact locations are required, then several issues must be considered. First, revealing exact locations may entice users to visit the plots to obtain additional information, thus artificially disturbing the sampling location and contributing to bias in inventory estimates. Second, plots may be located on private land, and while land owners usually permit access by inventory field crews, they are generally less receptive to access by non-inventory personnel. For some NFIs, these situations have potentially serious impacts. For the NFI of the USA, which must obtain owner permission to access plots on private land, unauthorized user visits to plot locations may jeopardize future access by inventory field crews. In addition, the program is prohibited by public law from revealing the exact locations of plots on private land. Thus, if exact plot locations are required for a user's analysis, policy constraints may prohibit the inventory program from accommodating the user's data request.

Additionally, users may require estimates for AOIs much smaller than those reported by an NFI. For some forest inventory variables, the precision of estimates based on plot data only for these small areas may be adequate. For example, Katila and Tomppo (2001) and Katila (2006) reported errors for mean volume of growing stock of 13% for forest holding as small as 100 ha. However, for many small AOIs of interest, the number of plots is not sufficient to obtain estimates with acceptable precision.

An alternative to plot-based estimation that addresses both the plot location security and small AOI problems is to construct maps of forest resources that are sufficiently unbiased that users may obtain estimates for small areas that are comparable to those obtained using dense field plot data alone, and more precise than those obtained using sparse field plot data. This alternative requires public access to the maps, probably via the Internet, and estimation algorithms that are sufficiently fast that the tolerance levels of users are not exceeded.

A complex component of map-based estimation is dealing with spatial correlation among observations on which the maps are based and among mapping unit predictions. The primary effect of spatial correlation is in the estimation of variances. Thus, if users are only interested in estimates of forest resources for their AOIs and have no requirement for measures of uncertainty, then the issue is not crucial. However, NFIs have traditionally provided not only forest resource estimates but also standard errors or confidence intervals associated with those estimates. For plot-based estimation, the small number of plots (at least relative to the number of pixels in an AOI) mitigates much of this complexity, particularly when the distances between plots exceed the range of spatial correlation. However, for map-based estimation, and for any remote sensing aided estimation method, the relatively small pixel separation distances virtually guarantee that spatial correlation cannot be ignored if defensible variance estimates are required. Thus, variance estimates for means, $\overline{Y}$, over multiple pixel AOIs must be expressed as,

$$\begin{aligned}\mathrm{Var}(\overline{Y}) &= \mathrm{Var}\left(\frac{1}{N}\sum_{i=1}^{N}\hat{y}_i\right) = \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathrm{Cov}(\hat{y}_i,\hat{y}_j)\\ &= \frac{1}{N^2}\left[\sum_{i=1}^{N}\mathrm{Var}(\hat{y}_i)+\sum_{i\neq}^{N}\sum_{j}^{N}\mathrm{Cov}(\hat{y}_i,\hat{y}_j)\right] \qquad (6)\end{aligned}$$

where N is the number of pixels in the AOI, and $\hat{y}_i$ and $\hat{y}_j$ are predictions for the ith and jth pixels, respectively. For $i\neq j$, it is not possible to assume $\mathrm{Cov}(\hat{y}_i,\hat{y}_j)=0$. Further, the expression for $\mathrm{Cov}(\hat{y}_i,\hat{y}_j)$ may be complex and typically must incorporate the spatial correlation among prediction errors. Estimation techniques for regression model-based approaches have been developed (McRoberts, 2006), and techniques for k-NN predictions have been developed as a joint effort between the Finnish and American NFIs (McRoberts et al., 2007-this issue). In addition, the large number of pixels for most AOIs means that substantial computing resources may be required, even for moderately sized AOIs.

5. Summary and conclusions

Remote sensing currently enhances NFIs in four primary ways: (1) providing faster and less expensive observation or measurement of some forest attributes, (2) increasing the precision of large area inventory estimates, often via stratified or weighted estimation, (3) providing inventory estimates with acceptable bias and precision for small areas for which sufficient field data are not available, and (4) producing forest thematic maps that can be used for purposes such as for timber production, procurement, and ecological studies. In addition, maps based on both field and remotely sensed data may be considered models of forests that can be used for applications such as simulating inventory sampling designs and comparing their efficiencies. Most inventory sampling designs are selected to support the objective of calculating estimates using field data. The efficiencies of inventories could be increased substantially if, in the planning phase, designers could rely on the availability of satellite based remote sensing data.

Both parametric and non-parametric estimation methods have been tested and applied in forest inventory applications. The keen interest in the non-parametric k-NN method is partly motivated by the desire to estimate simultaneously the large number of variables of interest.

Estimation of uncertainty in forest inventories is never a trivial task due to spatial correlation and trend-like changes in the variables of interest. The task becomes even more difficult when using multi-source data for which the spatial dependencies are particularly complex. Estimation of pixel-level bias and precision has received much attention in the literature. However, for multiple pixel AOIs, the methods have often been empirical and have been characterized by comparisons of estimates derived in two different ways, one of which produces less but known precision. Recent developments emphasizing derivations of analytic methods for error estimation are promising.

Two primary conclusions may be drawn from this brief review of current and future remote sensing applications in support of NFIs. First, satellite imagery has contributed greatly to the ability of NFIs to produce more timely, cost efficient, and precise inventory estimates and has greatly facilitated construction of the spatial products that are in increasing demand. Second, technologies that are now on the horizon have the potential to alter radically the ways in which trees are measured, estimates are produced, and products are delivered. The use of digital remote sensing data of different spatial and spectral resolutions can be expected to be an essential part of large area forest inventories. The future depends to a great degree on the availability of data and the development of statistically sound methods.

References

Breidt, F. J., & Opsomer, J. D. (2004). Endogenous post-stratification in surveys: Classifying with a sample fitted model. Preprint Series #02-23, Department of Statistics, Iowa State University.

Cochran, W. G. (1977). *Sampling techniques*, 3rd Ed. New York: Wiley.

Franco-Lopez, H., Ek, A. R., & Bauer, M. E. (2001). Estimation and mapping of forest stand density, volume, and cover type using the k-nearest neighbors method. *Remote Sensing of Environment*, *77*, 251−274.

Fransson, J. E. S., Walter, F., & Ulander, L. M. H. (2000). Estimation of forest parameters using CARABAS-II VHF SAR data. *IEEE Transactions on Geoscience and Remote Sensing*, *38*, 720−727.

Gjertsen, A. (2007). Accuracy of forest mapping based on Landsat TM data and a kNN method.*Remote Sensing of Environment*, *110*, 146−156 (this issue). doi:10.1016/j.rse.2006.08.018.

Halme, M., & Tomppo, E. (2001). Improving the accuracy of multisource forest inventory estimates by reducing plot location error — A multicriteria approach. *Remote Sensing of Environment*, *78*, 321−327.

Hassani, B. T., LeMay, V., Marshall, P. L., Temesegen, H., & Zumrawi, A. A. (2004). Regeneration imputation models for complex stands of southeaster British Columbia. *The Forestry Chronicle*, *80*, 271−278.

Henttonen, H. (1991). *VMI8:n Pohjois-Suomen otanta-asetelmien vertailu satelliittikuvatulkinnan avulla.* Comparisons of sampling designs in North Finland for NFI8. Manuscript. Finnish Forest Research Institute. (in Finnish).

Holmgren, J., Nilsson, M., & Olsson, H. (2003). Estimation of tree height and stem volume on plots using airborne laser scanning. *Forest Science*, *49*, 419−428.

Holström, H., & Fransson, J. E. (2003). Combining remotely sensed optical and radar data in *k*NN-estimation of forest variables. *Forest Science*, *49*, 409−418.

Homer, C., Huang, C., Yang, L., Wylie, B., & Coan, M. (2004). Development of a 2001 national landcover database for the United States. *Photogrammetric Engineering and Remote Sensing*, *70*(7), 829−840.

Hoppus, M. L., & Lister, A. J. (2003). A statistically valid method for using FIA plots to guide spectral class rejection in producing stratification maps. In R. E. McRoberts, G. A. Reams, P. C. Van Deusen, & J. W. Moser (Eds.), *Proceedings of the third annual forest inventory and analysis symposium. Traverse City, Michigan, USA, 17–19 October 2001. U.S. For. Serv. Gen. Tech. Rep. NC-230* (pp. 44−49).

Katila, M. (2006). Empirical errors of small area estimates from the multisource National Forest Inventory in Eastern Finland. *Silva Fennica*, *40*, 729−742.

Katila, M., & Tomppo, E. (2001). Selecting estimation parameters for the Finnish multisource National Forest Inventory. *Remote Sensing of Environment*, *76*, 16−32.

Kellndorfer, J., Walker, W., Pierce, L., Dobson, C., Fites, J. A., Hunsaker, C., et al. (2004). Vegetation height estimation from Shuttle Radar Topography Mission and National Elevation Datasets. *Remote Sensing of Environment*, *93*, 339−358.

Kendall, M. G., & Stewart, A. (1968). 2nd edition. *The advanced theory of statistics*, *Vol. 3* (pp. 259−260)NY: Hafner Publishing Co.

Koukal, T., Suppan, F., & Schneider, W. (2005). The impact of radiometric calibration on kNN predictions of forest attributes. In H. Olsson (Ed.), *Proceedings of ForestSat 2005. Borås, Sweden, 31 May–3 June 2005. Report 8a* (pp. 17−21).

Lefsky, M. A., Cohen, W. B., Acker, S. A., Parker, G. G., Spies, T. A., & Harding, D. (1999). Lidar remote sensing of the canopy structure and biophysical properties of Doublas-Fir western Hemlock forests. *Remote Sensing of Environment*, *70*, 339−361.

McInerney, D., Pekkarinen, A., & Haakana, M. (2005). Combining Landsat ETM+ with field data for Ireland's national forest inventory — A pilot study for Co. Clare. In H. Olsson (Ed.), *Proceedings of ForestSat 2005. Borås, Sweden, 31 May–3 June 2005.* (pp. 12−16).

MCPFE (1990, December 18). *Ministerial Conference for the Protection of Forests in Europe. (1990).* General declaration and resolutions Strasbourg, France.

McRoberts, R. E. (2006). A model-based approach to estimating forest area. *Remote Sensing of Environment*, *103*, 56−66.

McRoberts, R. E., Bechtold, W. A., Patterson, P. L., Scott, C. T., & Reams, G. A. (2005). The enhanced Forest Inventory and Analysis program of the USDA Forest Service: Historical perspective and announcement of statistical documentation. *Journal of Forestry*, *103*(6), 304−308.

McRoberts, R. E., Holden, G. R., Nelson, M. D., Liknes, G. C., & Gormanson, D. D. (2006). Using satellite imagery as ancillary data for increasing the precision of estimates for the Forest Inventory and Analysis program of the United States of America. *Canadian Journal of Forest Research*, *36*, 2968−2980.

McRoberts, R. E., Nelson, M. D., & Wendt, D. G. (2002). Stratified estimation of forest area using satellite imagery, inventory data, and the k-nearest neighbors technique. *Remote Sensing of Environment*, *82*, 457−468.

McRoberts, R. E., Tomppo, E. O., Finley, A. O., & Heikkinen, J. (2007). Estimating areal means and variances of forest attributes using the k-Nearest Neighbors technique and satellite imagery. *Remote Sensing of Environment*. doi: 10.1016/j.rse.2007.04.002

McRoberts, R. E., Wendt, D. G., & Nelson, M. D. (2002). Using a land cover classification based on satellite imagery to improve the precision of forest inventory area estimates. *Remote Sensing of Environment*, *81*, 36−44.

Means, J. A., Acker, S. A., Harding, D. J., Blair, J. B., Lefsky, M. A., Cohen, W. B., et al. (1999). Use of large-footprint scanning airborne lidar to estimate forest stand characteristics in the western Cascades of Oregon. *Remote Sensing of Environment*, *67*, 298−308.

Moeur, M., & Stage, A. R. (1995). Most similar neighbor: An improved sampling inference procedure for natural resource planning. *Forest Science*, *41*, 337−359.

Montréal Process Working Group. (2005). *The Montréal Process.* http://www.mpci.org/home_e.html (accessed 01 August 2005).

Muinonen, E., & Tokola, T. (1990). An application of remote sensing for communal forest inventory. In R. Sylvander & S. Poso (Eds.), *The usability of remote sensing for forest inventory and planning, proceedings from SNS/IUFRO workshop, Report 4* (pp. 35−42). Umeå, Sweden: Remote Sensing Laboratory, Swedish University of Agricultural Sciences.

Næsset, E. (1997). Estimating timber volume of forest stands using airborne laser scanner data. *Remote Sensing of Environment*, *61*, 246−253.

Nilsson, M. (1996). Estimation of tree heights and stand volume using an airborne lidar system. *Remote Sensing of Environment*, *56*, 1−7.

Nilsson, M., Folving, S., Kennedy, P., Puumalainen, J., Chirici, G., Corona, P., et al. (2003). Combining remote sensing and field data for deriving unbiased estimates of forest parameters over large regions. In P. Corona, M. Köhl, & M. Marchetti (Eds.), *Advances in forest inventory for sustainable forest* *management and biodiversity monitoring* (pp. 19−32). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Ohmann, J. L., & Gregory, M. J. (2002). Predictive mapping of forest composition and structure with direct gradient analysis and nearest neighbor imputation in coastal Oregon, U.S.A. *Canadian Journal of Forest Research*, *32*, 725−741.

Pakkala, T., Hanski, I., & Tomppo, E. (2002). Spatial ecology of the three-toed woodpecker in managed forest landscapes. In E. Korpilahti & T. Kuuluvainen (Eds.), *Disturbance dynamics in boreal forests: Defining the ecological basis of restoration and management of biodiversity. Silva Fennica*, *36(1)* (pp. 279−288).

Parker, R. C., & Evans, D. L. (2004). An application of LiDAR in a double-sample forest inventory. *Western Journal of Applied Forestry*, *19*, 95−101.

Parker, R. C., & Glass, P. A. (2004). High- versus low-density lidar in a double-sample forest inventory. *Southern Journal of Applied Forestry*, *28*(4), 205−210.

Poso, S., Halme, T., & Paananen, R. (1984). A method of estimating the stand characteristics of a forest component using satellite imagery. *Silva Fennica*, *18*, 261−292.

Poso, S., Paananen, R., & Simila, M. (1987). Forest inventory by compartments using satellite imagery. *Silva Fennica*, *21*, 69−94.

Reese, H., Nilsson, M., Pahlén, Hagner, O., Joyce, S., Tingelöf, U., et al. (2003). Countrywide estimates of forest variables using satellite data and field data from the National Forest Inventory. *Ambio*, *32*, 542−548.

Tokola, T. (2000). The influence of field sample data location on growing stock volume estimation in Landsat TM-based forest inventory in eastern Finland. *Remote Sensing of Environment*, *74*, 422−431.

Tokola, T., Pitkänen, Partinen, S., & Muionen, E. (1996). Point accuracy of a non-parametric method in estimation of forest characteristics with different satellite materials. *International Journal of Remote Sensing*, *17*, 2333−2351.

Tomppo, E. (1991). Satellite image-based National Forest Inventory of Finland. *Proceedings of the symposium on global and environmental monitoring, techniques and impacts, Victoria, British Columbia, Canada, 17–21 September 1990. International Archies of Photogrammetry and Remote Sensing*, *Vol. 28 (part 1–7)* (pp. 419−424).

Tomppo, E. (1996). Multi-source National Forest Inventory of Finland. In R. Päivinen, J. Vanclay, & S. Miina (Eds.), *New thrusts in forest inventory, proceedings of the Subject Group S4.02-00, Forest Resource Inventory and Monitoring, and Subject Group S4.12-00, Remote Sensing Technology, volume I. IUFRO XX World Congress, Tampere, Finland, 6–12 August 1995. EFI, EFI Proceedings*, *Vol. 7* (pp. 27−41).

Tomppo, E., & Halme, M. (2004). Using coarse scale forest variables as ancillary information and weighting of k-NN estimation: A genetic algorithm approach. *Remote Sensing of Environment*, *92*, 1−20.

Tomppo, E., Henttonen, H., & Tuomainen, T. (2001). Valtakunnan metsien 8. inventoinnin menetelmä ja tulokset metsäkeskuksittain Pohjois-Suomessa 1992–94 sekä tulokset Etelä-Suomessa 1986–92 ja koko maassa 1986–94. *Metsätieteen Aikakauskirja*, *1B/2001*, 99−248. (in Finnish).

Tomppo, E., Nilsson, M., Rosengren, M., Aalto, P., & Kennedy, P. (2002). Simultaneous use of Landsat-TM and IRS-1C WiFS data in estimating large area tree stem volume and aboveground biomass. *Remote Sensing of Environment*, *82*, 156−171.

Vogelmann, J. E., Howard, S. M., Yang, L., Larson, C. R., Wylie, B. K., & Van Driel, N. (2001). Completion of the 1990s National Land Cover Data Set for the conterminous United States from Landsat Thematic Mapper data and ancillary data sources. *Photogrammetric Engineering and Remote Sensing*, *67*, 650−662.

Wallerman, J. P., & Holmgren, J. (2007). Data capture for forest management planning using sample plot imputation based on spatial statistics, laser scanner, and satellite image data.*Remote Sensing of Environment*, *110*, 227−234 (this issue). doi:10.1016/j.rse.2007.02.028.

Woodhouse, I. H., & Hoekman, D. H. (2000). Radar modelling of coniferous forest using a tree growth model. *International Journal of Remote Sensing*, *21*, 1725−1737.